

REMARKS

In the September 17, 2003 Office Action, the Examiner noted that claims 1-21 were pending in the application; rejected claims 1-4, 6, 10, 11, 14-16 and 18-31 under 35 U.S.C. § 102(e); and rejected claims 5, 7-9, 12, 13 and 17 under 35 U.S.C. § 103. In rejecting the claims, U.S. Patents 6,078,913 to Aoki et al. and 6,212,517 to Sato et al. (References A and B, respectively) were cited. Claims 1-21 remain in the case. The Examiner's rejections are traversed below.

The Invention

The present invention is directed to collecting documents accessible via a network that are linked to each other, e.g., by hypertext in, e.g., hypertext markup language (HTML) on what is commonly termed the Internet, or on one or more Intranets. The collection process starts with an initial document group that is generated for documents found within what the application refers to as a "community". As described on page 26, a community can be "an in-house site, an industrial site (business community), ... a user group in a network of a specified topic, etc." (page 6, lines 11-14) where an in-house site may be an Intranet having "Intranet portal in an enterprise referenced as a corporate portal (also referenced as an EIP (Enterprise Information Portal))" (page 6, lines 19-22). An industrial site may be an "Extranet formed by a plurality of corporation systems" (page 26, lines 16-17). The initial document group maybe obtained in any known way, such as by frequency of reference among all documents within the community or identified as important to the community, e.g., by users as a whole, or a subset of users, such as supervisors, officers, guides, etc.

Using the initial document group as a collected document groups, a collection process starts with other documents within the community that are referenced by documents in the collected document group, e.g., by a hyperlink or URL in one of the documents in the collected document group pointing to a prospective document within the community. After a predetermined number of documents have been collected from inside the community, the collection process continues with documents both inside and outside the community. Prospective documents considered for collection may be ranked according to position in the network co-reference scores, keywords based on the referencing expressions used in the collected documents and the number of occurrences of the referencing expressions, or information about the field of the documents.

The Prior Art**U.S. Patents 6,212,517 to Sato et al.**

The Sato et al. patent is directed to the selection of related keywords extracted from documents and retrieving documents based on the keywords. As illustrated in Fig. 4, the system disclosed by Sato et al. works on a defined database of documents. All of the words in the documents are processed to obtain statistics indicating how many times each word appears (see the local statistic tables 220 in Fig. 4). Nothing has been found in Sato et al. suggesting any kind of link, referencing, or pointing from one document to another.

U.S. Patent 6,078,903 to Aoki et al.

The Aoki et al. patent is directed to a document retrieval apparatus in which the results of a search in response to a query are presented to a user using a hierarchical tree structure in which documents with similar contents are located closely to each other. To enable the generation of the results in this manner, contents database 123 (Fig. 1B) stores document index information including the information required to determine the closeness of retrieved documents. No description of the process of generating the contents database has been found in Aoki et al.

Rejections under 35 U.S.C. § 102(e)

In item 2 on pages 2-6 of the Office Action, claims 1-4, 6, 10, 11, 14-16 and 18-21 were rejected under 35 U.S.C. § 102(e) as anticipated by Sato et al. As described above, Sato et al. does not disclose "collecting documents linked to each other from a network by crawling the network" (e.g., claim 1, 1-2). Rather, Sato et al. performs all operations on a predefined set of documents of a number small enough that every single word can be indexed as required by the method taught by Sato et al. The present invention, on the other hand, is designed for use with the Internet where the documents available are constantly changing and are so large in number that indexing every single word of every single document is impractical.

In addition, nothing has been cited or found in Sato et al. corresponding to "community" as that term is defined in the application. In rejecting claim 1, the document database 70 in Fig. 2 is cited as corresponding to the community. However, if it is assumed that the document database is the community, nothing has been cited or found of anything outside the community on which the second collecting operation recited in claim 1 can be performed. Furthermore, the Office Action referenced server 1a in Fig. 14 in the paraphrase of the second operation of claim 1. However, server 1a is described as representing "the rest of the document retrieval system"

(column 12, line 61) where the user interface 400 is represented by client terminal 7 in Fig. 14. It is unclear how the server 1a could perform the second collection operation recited in claim 1.

Thus, there is no suggestion in Sato et al. of performing a collection process in two operations, first "collecting documents ... [until] a predetermined value [is reached] from inside a community" (claim 1, lines 3-4) and then "collecting documents from inside and outside the community" (claim 1, line 5). Since claims 2-4, 10, 11, 14 and 15 depend from claim 1, it is submitted that claims 1-4, 10, 11, 14 and 15 patentably distinguish over Sato et al.

In rejecting claim 6, the only reference to anything disclosed by Sato et al. was table 250 in Fig. 8, the generation of which is described in columns 5 and 6. Nothing was cited or has been found in Sato et al. regarding whether table 250 could be considered a positive or negative sample group or how it might be used in determining whether a document is to be collected. Therefore, it is submitted that anticipation of claim 6 has not been shown.

In rejecting claim 16, the only part of Sato et al. that was cited was the Abstract. As discussed above with respect to claim 1, nothing has been found anywhere in Sato et al., including the Abstract, related to retrieving documents "separately from inside and outside the community" (claim 16, lines 5-6). Furthermore, since nothing has been found in Sato et al. corresponding to a "community", there is no suggestion of any "information indicating a significance level for the community" (claim 16, last two lines). For the above reasons, it is submitted that claim 16 patentably distinguishes over Sato et al.

In rejecting claims 18 and 20, the only thing cited from Sato et al. was the results table illustrated in Fig. 8. Therefore, for the same reasons as discussed above with respect to claim 6, it is submitted that there is a lack of evidence of anticipation of claims 18 and 20.

In rejecting claim 19, the following portions of Sato et al. were cited: column 7, lines 43-52; column 6, lines 33-48; and server 1a in Fig. 14. The lack of relevance of server 1a was discussed above with respect to claim 1. The text cited in rejecting claim 17 describes the sorting of records and display of, e.g., 10 records with the greatest importance, along with statistical information. Also, storage of keywords and degree of importance as a history of the user is mentioned. The cited portion of column 7 does not appear to have any relevance, since it simply notes that high-precision document retrieval insuring retrieval of at least one document is provided by the system and that modifications are possible. As discussed above with respect to claim 1, nothing has been found in these portions or any other portion of Sato et al. regarding first collecting documents from a community and then "collecting documents from inside and outside the community" (claim 19, line 6) after a predetermined number of documents have been

collected from inside the community. Therefore, it is submitted that claim 19 patentably distinguishes over Sato et al. for the reasons discussed above with respect to claim 1.

In rejecting claim 21, nothing was cited in Sato et al. As a result, it is impossible to explain why claim 21 distinguishes over Sato et al., other than to note that limitations similar to those recited in claim 1 are recited in claim 21 and therefore, claim 21 patentably distinguishes over Sato et al. for the reasons discussed above with respect to claim 1

Rejections under 35 U.S.C. § 103

In item 4 on pages 6-8 of the Office Action, claims 7-9, 12 and 13 were rejected under 35 U.S.C. § 103(a) as unpatentable over Sato et al. Claims 12 and 13 depend from claim 1 and claims 7-9 depend from claim 6. Therefore, it is submitted that claims 7-9, 12 and 13 patentably distinguish over Sato et al. for the reasons discussed above, even if the additional limitations recited in claims 7-9, 12 and 13 are either disclosed by Sato et al. or would be obvious to one of ordinary skill in the art.

In item 5 on pages 9-10 of the Office Action, claims 5 and 17 were rejected under 35 U.S.C. § 103(a) as unpatentable over Sato et al. in view of Aoki et al. Claim 5 depends from claim 1 and therefore, it is submitted that claim 5 patentably distinguishes over Sato et al. taken alone for the reasons discussed above with respect to claim 1. As described above, Aoki et al. does not disclose anything that would overcome the deficiencies of Sato et al. discussed above with respect to claim 1. Therefore, it is submitted that claim 5 patentably distinguishes over the combination of Sato et al. and Aoki et al.

Claim 17 recites "determining whether or not the prospect is in a community in the network" (claim 17, lines 5-6) and collecting "the prospect from inside and outside the community after collecting documents larger in number than a predetermined value from inside the community" (claim 17, last two lines). Since Aoki et al. does not add anything to the teachings of Sato et al. regarding operations inside and outside of a community, as that term is defined in the specification, it is submitted that claim 17 patentably distinguishes over Sato et al. in view of Aoki et al. for the reasons discussed above with respect to claim 1.

Summary

It is submitted that the references cited by the Examiner do not teach or suggest the features of the present claimed invention. Thus, it is submitted that claims 1-21 are in a condition suitable for allowance. Reconsideration of the claims and an early Notice of Allowance are earnestly solicited.

Finally, if there are any formal matters remaining after this response, the Examiner is requested to telephone the undersigned to attend to these matters.

If there are any additional fees associated with filing of this Amendment, please charge the same to our Deposit Account No. 19-3935.

Respectfully submitted,

STAAS & HALSEY LLP

Date: 12/17/03

By: Richard A. Gollhofer
Richard A. Gollhofer.
Registration No. 31,106

1201 New York Avenue, NW, Suite 700
Washington, D.C. 20005
Telephone: (202) 434-1500
Facsimile: (202) 434-1501

CERTIFICATE UNDER 37 CFR 1.8(a)
I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to Commissioner for Patents, PO Box 1450, Alexandria, VA 22313-1450
on 12/17/03
STAAS & HALSEY
By: Richard A. Gollhofer
Date: 12/17/03